PRE2 - CLASS 6

EXERCICE 3 – CLASS 3

Let $\{v_1, \dots, v_n\}$ be n non-zero vectors such that $v_i \cdot v_j = 0 \quad i \neq j$

Show that $\{v_1, \dots, v_n\}$ are linearly independent.

Let $\lambda_1, \dots, \lambda_n$ n scalars such that $\lambda_1 v_1 + \lambda_2 v_2 + \dots + \lambda_n v_n = 0$

We need to show that all $\lambda_i = 0$.

Let's take the dot product between $\lambda_1 v_1 + \dots + \lambda_n v_n$ and v_1:

$$\Big(\underbrace{\lambda_1 v_1 + \dots + \lambda_n v_n}_{=0}\Big) \cdot v_1 = \lambda_1 v_1 \cdot v_1 + \lambda_2 \underbrace{v_1 \cdot v_2}_{=0} + \dots + \lambda_n \underbrace{v_n \cdot v_1}_{=0} = \lambda_1 v_1 \cdot v_1$$
$$= \lambda_1 \underbrace{\|v_1\|^2}_{\neq 0}$$

Then $\lambda_1 \underbrace{\|v_1\|^2}_{\neq 0} = 0 \implies \lambda_1 = 0$

By repeating this procedure with each v_i, we find that $\lambda_i = 0 \quad \forall i$.

EXERCICE 2 - CLASS 4.

We start from a norm on $\mathbb{R}^n$, (on vectors), $\|\cdot\|$ $\left[\text{e.g } \|x\| = \left(\sum_{i=1}^{n} x_i^2\right)^{1/2}\right.$

$\left.\|\cdot\|_\infty, \|\ \|_1\right]$

We define $|||A||| = \sup_{x \neq 0} \frac{\|Ax\|}{\|x\|} = \sup_{\|x\|=1} \|Ax\|$

$x \neq 0 \quad \overset{f}{\longmapsto} \quad \frac{\|Ax\|}{\|x\|}$

$\forall x \quad f(x) \leq \underset{x}{\sup / \max}\, f(x)$

1/ We want to show that $\|Ax\| \leq |||A|||\, \|x\| \quad \forall x$

Since $|||A||| = \sup_{x \neq 0} \frac{\|Ax\|}{\|x\|} \Rightarrow \forall x \neq 0 \quad \frac{\|Ax\|}{\|x\|} \leq |||A|||$

Then $\forall x \neq 0 \quad \|Ax\| \leq |||A|||\, \|x\|$

If $x = 0$ we also have that $\|Ax\| = 0 \leq \underbrace{|||A|||\, \|x\|}_{=0}$ ($\|x\| = 0$)

Then we have $\forall x, \ \|Ax\| \leq |||A|||\, \|x\|$

2) We want to prove that $|||AB||| \leq |||A||| \; |||B|||$

From 1 we know that $\|ABx\| \leq |||A||| \, \|Bx\| \leq |||A||| \, |||B||| \, \|x\|$

For $x \neq 0$, we have $\|ABx\| \leq |||A||| \, |||B||| \, \|x\|$

$$\Rightarrow \quad \frac{\|ABx\|}{\|x\|} \leq |||A||| \, |||B||| \qquad \forall x \neq 0$$

$$\Rightarrow \quad \underbrace{\sup_{x \neq 0} \frac{\|ABx\|}{\|x\|}}_{= |||AB|||} \leq |||A||| \, |||B|||$$

We have shown that $|||AB||| \leq |||A||| \, |||B|||$

3) Given a symmetric matrix A, and the Euclidean norm, we want to show that

$$|||A||| = \max_{i=1,\dots,n} |\lambda_i|$$ where λ_i are the eigenvalues of A.

A is symmetric, it has n eigenvalues $\lambda_1, \dots, \lambda_n$

Let u be the eigenvector associated to λ such that $|\lambda| = \max_{i=1,\dots,n} |\lambda_i|$

Then $Au = \lambda u$ by definition.

Because $|||A||| = \sup_{x \neq 0} \frac{\|Ax\|}{\|x\|}$, then $\frac{\|Au\|}{\|u\|} \leq \sup_{x \neq 0} \frac{\|Ax\|}{\|x\|} = |||A|||$

$$\frac{\|Au\|}{\|u\|} = \frac{\|\lambda u\|}{\|u\|} = |\lambda| \frac{\|u\|}{\|u\|} = |\lambda|$$

$\frac{\|Au\|}{\|u\|} = |\lambda|$

We have $|\lambda| \leq |||A|||$

A is symmetric, then from the spectral theorem: $A = VDV^T$ with V: orthogonal.

$$\|Ax\| = \|VDV^Tx\| \underset{\text{since } V \text{ is orthogonal}}{=} \|D\overbrace{V^Tx}^{=y}\| = \sqrt{\sum_{i=1}^{n} (\lambda_i y_i)^2}$$

[Because $Dy = \begin{pmatrix} \lambda_1 y_1 \\ \vdots \\ \lambda_n y_n \end{pmatrix}$

$$= \sqrt{\sum_{i=1}^{n} \underbrace{\lambda_i^2 y_i^2}_{\leq \lambda^2 y_i^2}} \leq \sqrt{\sum_{i=1}^{n} \lambda^2 y_i^2} = \sqrt{\lambda^2 \underbrace{\sum_{i=1}^{n} y_i^2}_{\|y\|^2}} = |\lambda| \|y\| = |\lambda| \|\underbrace{V^Tx}_{\substack{=\|x\| \\ V \text{ orthogonal}}}\| = |\lambda| \|x\|$$

We have $\lambda_i^2 \leq \lambda^2 \ \forall i$

We have $\|Ax\| \leq |\lambda| \|x\|$

$$|\lambda| = \max_{i=1,\dots,n} |\lambda_i| \implies |\lambda_i| \leq |\lambda|$$

$$\implies \underbrace{|\lambda_i|^2}_{\lambda_i^2} \leq \underbrace{|\lambda|^2}_{\lambda^2}$$

How do we prove that V^T is orthogonal ?

V is orthogonal $\quad V^T V = V V^T = Id$

To prove that V^T is orthogonal $\wedge$ $(V^T)^T V^T = Id$ and $V^T (V^T)^T = Id$

we need

1) $(V^T)^T V^T = V V^T = Id$ because of

2) $V^T (V^T)^T = V^T V = Id$ because of

We have shown that $\|Ax\| \leq |\lambda| \, \|x\|$

$$\Rightarrow \quad \frac{\|Ax\|}{\|x\|} \leq |\lambda|$$

$$\Rightarrow \quad \underbrace{\sup_{x \neq 0} \frac{\|Ax\|}{\|x\|}}_{|||A|||} \leq |\lambda|$$

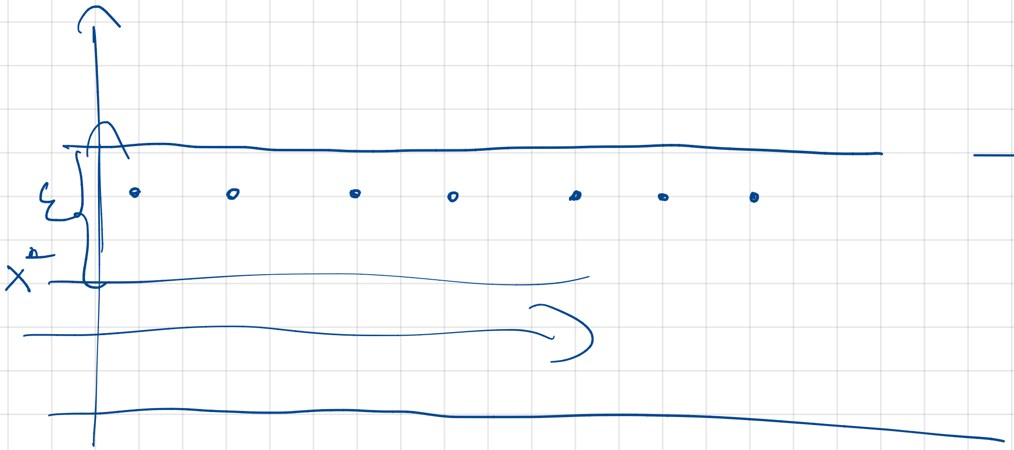
We have shown that $|||A||| \leq |\lambda|$ and $|\lambda| \leq |||A|||$

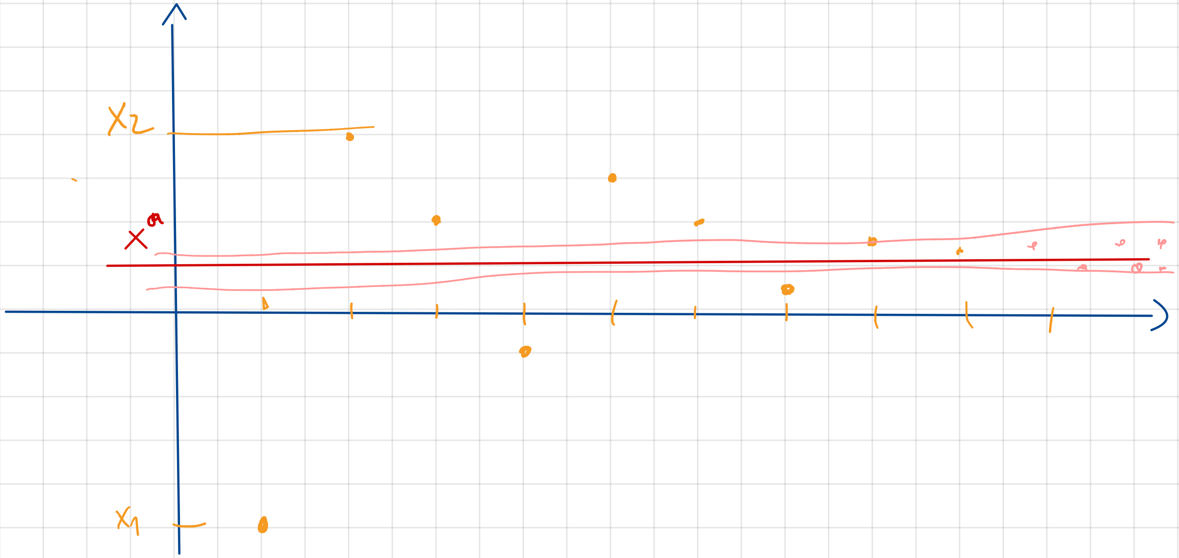
This implies that $|||A||| = |\lambda| = \max\limits_{i=1,\dots,n} |\lambda_i|$

CONVERGENCE OF A SEQUENCE

Given a sequence of real numbers $\{x_n, n \in \mathbb{N}\}$ we say that x_n converges to x^*, denoted as $\lim_{n \to +\infty} x_n = x^*$ if

$$\forall \varepsilon > 0, \quad \exists N, \quad \forall n \geq N \quad |x_n - x^*| < \varepsilon$$





DIVERGENCE TO $+\infty$:

$\{x_n, n \in \mathbb{N}\}$ diverges to $+\infty$ if

$$\forall A > 0, \quad \exists N, \quad \forall n \geq N \quad x_n \geq A$$

Examples: $\quad x_n = \frac{1}{n} \qquad x_n = \frac{1}{n^2} \qquad x_n = n \qquad x_n = \frac{(-1)^n}{n} \qquad x_n = (-1)^n$

Limits of those sequences. $\qquad n \geq 1 \quad n \in \mathbb{N}$

$$\lim_{n \to +\infty} \frac{1}{n} = 0$$

$n = 1 \longrightarrow \frac{1}{n} = 1$

$n = 2 \longrightarrow \frac{1}{n} = \frac{1}{2}$

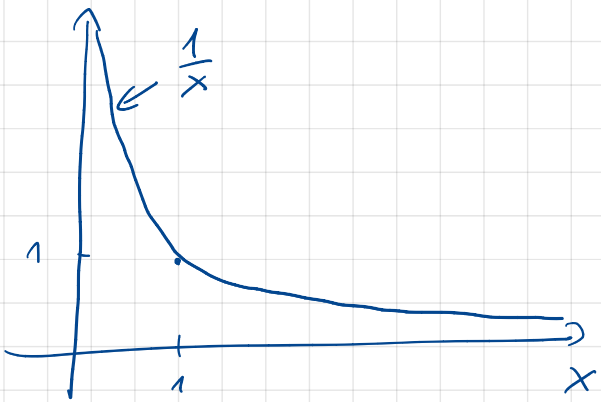
$n = 3 \longrightarrow \frac{1}{3}$

$\vdots$

$n = 100 \longrightarrow 0{,}01$

$\vdots$

$$\lim_{n \to +\infty} \frac{1}{n} = 0$$



$$\lim_{n \to +\infty} \frac{1}{n^2} = 0 \qquad \lim_{n \to +\infty} \overset{n}{\overset{\shortparallel}{x_n}} = +\infty \qquad \lim_{n \to +\infty} \frac{(-1)^n}{n} = 0,$$

$x_n = (-1)^n$

$$\left.\begin{aligned} x_1 &= -1 \\ x_2 &= 1 \\ x_3 &= -1 \\ x_4 &= 1 \end{aligned}\right\} \longrightarrow \text{no limit.}$$

How do we generalize the definition of convergence to a sequence of vectors?

Given a norm on $\mathbb{R}^n$, $\| \cdot \|$, we say that a sequence of vectors $\{x_n \in \mathbb{R}^n, n \in \mathbb{N}\}$ converges to $x^* \in \mathbb{R}^n$ if:

$$\forall \varepsilon > 0, \quad \exists N, \quad \forall n \geq N, \quad \| x_n - x^* \| \leq \varepsilon$$

We can generalize the definition to matrices, using a norm on matrices.

BIG O and little o :

$$x_n = O(\overset{y_n}{n^2})$$ what does it mean ?

$\exists\, C$, $\exists\, N$ such $n \geq N$

$$|x_n| \leq C n^2$$

$$|x_n| \leq C y_n$$

$$x_n = o(y_n)$$

$\forall\, \varepsilon > 0$, $\exists\, N$, $\forall n \geq N$ $\quad |x_n| \leq \varepsilon y_n$

ORDER THE FOLLOWING IN TERMS OF O

$\exp(n)$, n, $\log(n)$, $n \log n$

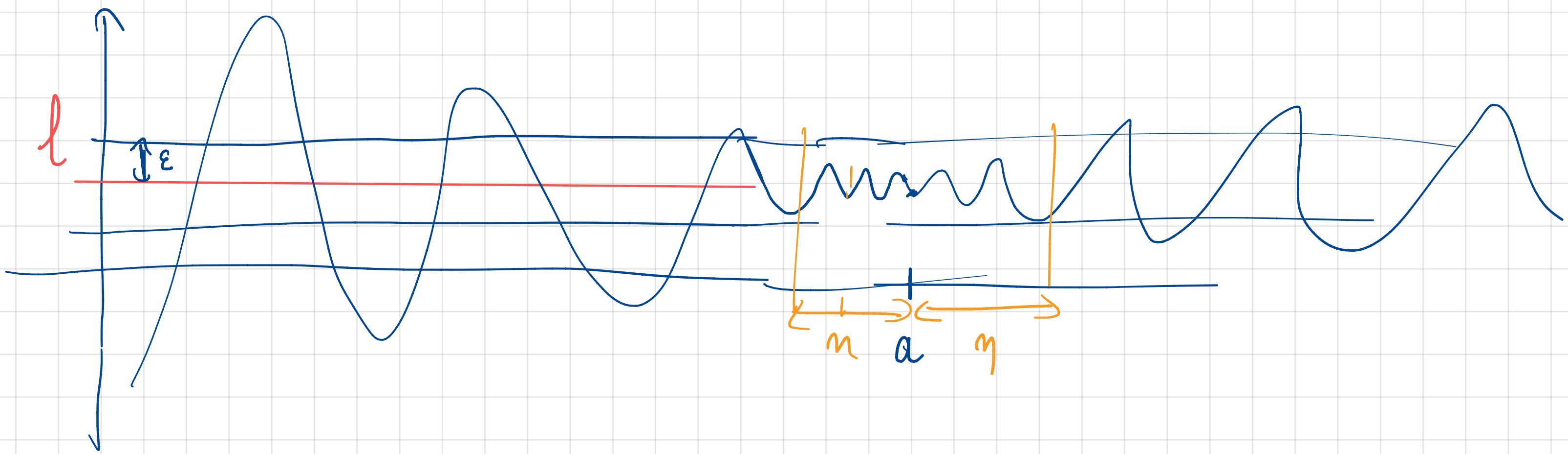
$\log(n) = O(n) \qquad n = O(n \log n) \qquad n \log(n) = o(\exp(n))$

GENERALIZATION TO FUNCTIONS

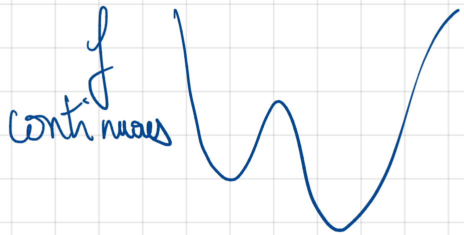
Let $f : \mathbb{R} \to \mathbb{R}$, we say that f converges to ℓ when x goes to a

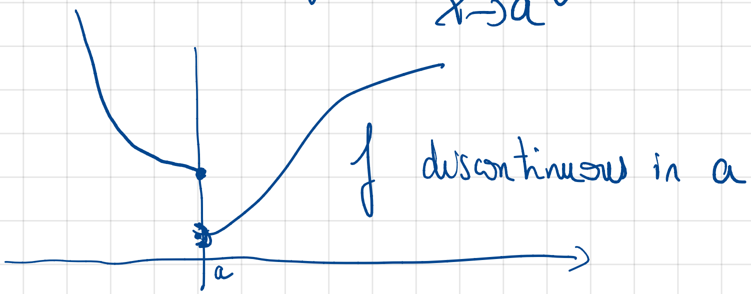
denoted $\lim\limits_{x \to a} f(x) = \ell$

$$\forall \varepsilon > 0,\ \exists \eta,\ \forall x\ |x - a| < \eta \Rightarrow |f(x) - \ell| < \varepsilon$$



A function f is continuous in a, if $\lim\limits_{x \to a} f(x) = f(a)$





$f(x) \underset{x \to a}{=} o(g(x))$ if $\forall \varepsilon > 0, \exists \eta > 0, \forall x \ |x-a| < \eta \Rightarrow |f(x)| \leq \varepsilon \, |g(x)|$

$\left(\lim_{x \to a} \frac{f(x)}{g(x)} = 0 \right)$

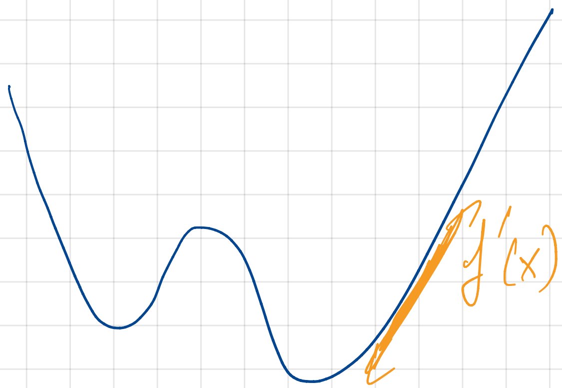
$f(x) \underset{x \to a}{=} O(g(x))$, $\exists C > 0, \exists \eta > 0, \forall x, |x-a| < \eta$

$|f(x)| \leq C |g(x)|$

DERIVATIVE

Let $f: \mathbb{R} \to \mathbb{R}$, then f is derivable in x

if $\lim_{h \to 0} \frac{f(x+h) - f(x)}{h}$ exists, the limit is called derivative and denoted $f'(x)$



$f'(x)$ is the slope of f in x.

If f is derivable in x then $\forall h \in \mathbb{R}$ $\underbrace{f(x+h) = f(x) + h f'(x) + o(|h|)}$

TAYLOR FORMULA, order 1

Let us take a function $f : \begin{cases} \mathbb{R}^n \longrightarrow \mathbb{R} \\ (x_1, \dots, x_n) \longrightarrow f(x_1, \dots, x_n) \end{cases}$

how do we generalize the derivative for such a function ?

PARTIAL DERIVATIVE :

$$\frac{\partial f}{\partial x_i} = \lim_{h \to 0} \frac{f(x_1, \dots, x_{i-1}, x_i + h, x_{i+1}, \dots, x_n) - f(x)}{h}$$

Partial derivative of f with respect to x_i

We define the vector $Df(x) = \begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{pmatrix}$ which is called the gradient of f in x.

<u>Taylor formula</u> : If f is differentiable then

$$\forall h \in \mathbb{R}^n \qquad f(x+h) = f(x) + Df(x) \cdot h + o(\|h\|)$$

dot product.

~~Com~~ EXERCICE : Compute Df where

$$f_1(x_1, x_2) = x_1^2 + x_2^2$$

$$f_2(x_1, x_2) = \frac{1}{2} x^T D x \qquad D = \begin{pmatrix} 9 & 0 \\ 0 & 1 \end{pmatrix}$$

<u>ANSWER</u>: $Df_1 = \begin{pmatrix} 2x_1 \\ 2x_2 \end{pmatrix}$; $Df_2 = \begin{pmatrix} 9x_1 \\ x_2 \end{pmatrix}$

CHAIN RULE

$f : \mathbb{R} \longrightarrow \mathbb{R}$, $g : \mathbb{R} \longrightarrow \mathbb{R}$

$$(f \circ g)'(x) = \underbrace{(f' \circ g)(x) \times g'(x)}_{f'(g(x)) \times g'(x)}$$

Let $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$, we say that f is differentiable in $x \in \mathbb{R}^n$ if there exist a linear mapping $Df(x) : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ such that

$$\forall h \in \mathbb{R}^n, \quad f(x+h) = f(x) + \overbrace{Df(x)(h)}^{L(h)} + o(\|h\|)$$

$Df(x)$ is called the differential of f in x

If $f : \mathbb{R}^n \to \mathbb{R}$, then $Df(x)(h) = \nabla f(x) \cdot h$

$$f : \begin{cases} \mathbb{R}^n \longrightarrow \mathbb{R}^n \\ x \longrightarrow Ax \end{cases}$$

where A $n \times n$ matrix.

$$Df(x) : \begin{cases} \mathbb{R}^n \longrightarrow \mathbb{R}^n \\ h \longrightarrow Ax \end{cases}$$

[CHECK AT HOME]

CHAIN RULE:

$$\boxed{D(f \circ g)(x) = Df(g(x)) \circ Dg(x)}$$

$\longrightarrow Df(g(x))\left(Dg(x)\right)$

Let us consider $f: x \in \mathbb{R}^n \longrightarrow Ax$ for A a $n \times n$ matrix.

To find out wheter f is differentiable we have to see if we can write $f(x+h) = f(x) + \underbrace{\quad\quad}_{\text{something linear in } h} + o(\|h\|)$

$$f(x+h) = A(x+h) = Ax + Ah$$
$$= f(x) + Ah$$

Since $h \longmapsto Ah$ is linear in h

Then $Df(x)(h) = Ah$

In general this is more complex. , A symmetric.

Compute differential of $f(x) = \frac{1}{2} x^T A x$